UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

June 5, 2000

OFFICE OF
THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

Note to the Reader:

The attached draft report is a draft report of a Joint Subcommittee of the Science Advisory Board (SAB) and the Scientific Advisory Panel (SAP). The draft is still undergoing final internal review, however, in its present form, it represents the consensus position of the panel involved in the review. The Chair of the SAP will participate with the SAB Executive Committee in the final review. Once approved as final, the report will be transmitted to the EPA Administrator and will become available to the interested public as a final report.

This draft has been released for general information to members of the interested public and to EPA staff. This is consistent with our policy of releasing draft materials only when the Committee involved is comfortable that the document is sufficiently complete to provide useful information to the reader. The reader should remember that this is an unapproved working draft and that the document should not be used to represent official EPA, SAP, or SAB views or advice. Draft documents at this stage of the process often undergo significant revisions before the final version is approved and published.

The SAP/SAB Joint Subcommittee is not soliciting comments on the advice contained herein. However, as a courtesy to the EPA Program Office which is the subject of the review, we have asked them to respond to the issues listed below. Consistent with our policy on this matter, the SAP/SAB Subcommittee is not obligated to address any responses which it receives.

1. Has the Committee adequately responded to the questions posed in the Charge?
2. Are any statements or responses made in the draft unclear?
3. Are there any technical errors?

For further information or to respond to the questions above, please contact:

Sam Rondberg, Designated Federal Officer
Science Advisory Board (1400A)
US Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460
(301) 812-2560 Fax: (410) 286-2689
E-Mail: samuelr717@aol.com

EXECUTIVE COMMITTEE REVIEW DRAFT

REPORT OF THE JOINT SAB/SAP SUBCOMMITTEE ON DATA FROM TESTING OF HUMAN SUBJECTS

May 31, 2000

DRAFT FOR REVIEW ONLY -- DO NOT CITE OR QUOTE

EPA-SAB-EC-00-00X

Honorable Carol Browner
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460

Subject: Consideration of issues relating to EPA's use of data derived from the testing of human subjects

Dear Ms. Browner:

A Joint Subcommittee of the Science Advisory Board (SAB) and the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA) Scientific Advisory Panel (SAP) convened in a public meeting on December 10-11, 1998. The purpose of the meeting was to provide advice and comment to the Environmental Protection Agency (EPA) on issues related to data derived from testing on human subjects, particularly the use of human data for making pesticide registration decisions. Both scientific and ethical questions have been raised about the data, to include the manner in which they were developed and how or whether these data should be used in the decision making process. A draft report was generated based on the presentations and discussions at this meeting. However, a significant subset of the Subcommittee had reservations about the content of some sections of the report. Therefore, a second meeting of the SAB/SAP Joint Subcommittee was convened on November 30, 1999 to permit further deliberations for the purpose of resolving and bringing to closure differences of opinion within the Committee.

The specific Charge to the Subcommittee addressed the value of human studies; factors for consideration when determining what constitutes an appropriate human study for use in environmental decision-making, when making a judgment on what constitutes an ethically appropriate human study, and when determining if a study is appropriate (or inappropriate) for use; the risks and benefits to subjects and society; and issues in determining if studies are in compliance with accepted guidelines (the complete Charge will be found in section 2.2 of this report).

Section 3 of the report addresses each element of the Charge, and provides many specific recommendations to the EPA. The Subcommittee found, however, that its most significant findings could be best expressed outside the constraints of specific Charge issues. These findings are:

a) All research involving humans should require prior review by an Institutional Review Board (IRB).

b) The structure, function, and activities of both the Agency's IRBs and external IRBs of entities submitting data should be under active and aggressive scrutiny by EPA, with adequate staff and financial resources provided to carry out this mission. EPA should establish an internal ethics review organization to perform this function, staffed by full-time individuals whose duties address exclusively compliance oversight. The review organization should also provide an institutional focus for continuous close liaison on ethical matters with other federal agencies.

c) The Subcommittee believes that intentional administration of pesticides to human subjects testing is acceptable, subject to limitations described as ranging from "rigorous" to "severe." Those Members supporting such testing feel that the information sought must not be available via other sources (e.g., animal studies and models, or the study of incidental exposures), and the information expected to be gained must promise reasonable benefits to the individual or society at large. Studies should be appropriately designed to address the stated objective, and have sufficient statistical power to provide an unambiguous answer to the question under investigation.

d) In no case should developing humans (i.e., the fetus, infant, and young children) be exposed to neurotoxic chemicals. There are currently too many unknown dangers to justify such studies, even under the most extraordinary circumstances.

e) The EPA should take whatever administrative action is necessary to extend the protections of 40 CFR Part 26 to all human research activities whose results will be submitted to the Agency.

f) Some of the Subcommittee Members that accepted the use of human volunteer testing of pesticides identified certain situations in which such testing would or would not be appropriate:

1) It would not be appropriate to conduct human volunteer testing when adequate

human data are already available.

2) Human volunteer studies would not be appropriate for pesticides in use today when data of equal quality can be obtained from field exposure studies.

3) Human volunteer studies could be appropriate when there are significant data gaps and such studies would provide a more accurate risk assessment.

4) Human volunteer studies could be appropriate for pesticides, which are not yet on the market, i.e. new pesticides.

5) Given the significance of statistical considerations in regard to human study design, we encourage the Agency to organize a workshop to deal specifically with this issue.

We appreciate the opportunity to review these issues, and look forward to your response.

Sincerely,

Dr. Morton Lippmann, Interim Chair
Science Advisory Board

Dr. Ronald Kendall, Co-Chair
Data from the Testing of Human Subjects Subcommittee

Dr. Mark Utell, Co-Chair
Data from the Testing of Human Subjects Subcommittee

ENCLOSURE

NOTICE

This report has been written as part of the activities of the Science Advisory Board, a public advisory group providing extramural scientific information and advice to the Administrator and other officials of the Environmental Protection Agency. The Board is structured to provide balanced, expert assessment of scientific matters related to problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the Environmental Protection Agency, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use.

ABSTRACT

TO BE SUPPLIED

ROSTER
SAB/SAP JOINT SUBCOMMITTEE ON DATA
FROM TESTING HUMAN SUBJECTS
November 30, 1999

CO-CHAIRS

Dr. Ronald J. Kendall, Director and Professor, The Institute of Environmental and Human Health, Texas Tech University, Health Sciences Center, Lubbock, TX

Dr. Mark Utell, Professor of Medicine and Environmental Medicine, University of Rochester Medical Center, Rochester, NY

MEMBERS AND CONSULTANTS

Dr. Nancy Fiedler, Environmental and Occupational Health Sciences Institute, Dept. of Environmental & Community Medicine, Piscataway, NJ

Dr. Samuel Gorovitz, Professor of Philosophy and of Public Administration, Syracuse University, Syracuse, NY.

Dr. Jeffrey P. Kahn, Director, Center for Bioethics, University of Minnesota, Minneapolis MN

Dr. Ernest E. McConnell, ToxPath, Raleigh, North Carolina

Dr. Herb Needleman, University of Pittsburgh, School of Medicine, Pittsburgh, PA

Dr. J. Routt Reigart, Director, General Pediatrics, Medical University of South Carolina, Charleston SC

Dr. Bernard Weiss, Professor of Environmental Medicine and Pediatrics, University of Rochester School of Medicine and Dentistry, Rochester, NY

FEDERAL EXPERTS

Dr. Joseph J. DeGeorge, Associate Director for Pharmacology and Toxicology, Office of Review Management, Center for Drug Evaluation and Research, FDA, Rockville, MD

Dr. Gary B. Ellis, Director, Office for Protection from Research Risks, National Institutes of Health, Rockville, MD

Dr. Eric M. Meslin, Executive Director, National Bioethics Advisory Commission, Rockville MD

Dr. Chris Portier, National Institute of Environmental Health Sciences, Research Triangle Park, NC

CO-DESIGNATED FEDERAL OFFICIALS

Mr. Larry Dorsey, Executive Secretary, FIFRA Scientific Advisory Panel, Office of Pesticide Programs, Environmental Protection Agency, 401 M Street, SW (H7509C), Washington, D.C. 20460

Mr. Samuel Rondberg (1400), Designated Federal Official, Science Advisory Board, Environmental Protection Agency, 401 M Street, SW, Washington, D.C. 20460

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

The Joint Science Advisory Board/Scientific Advisory Panel (SAB/SAP) Data from Testing on Human Subjects Subcommittee (DTHSS) first met on December 10-11, 1998, in Arlington VA to consider a series of issues raised by the EPA Office of Pesticides Programs concerning the acquisition and use of data generated by testing human subjects. The Charge addressed the value of human studies; factors for consideration when determining what constitutes an appropriate human study for use in environmental decision-making, when making a judgment on what constitutes an ethically appropriate human study; and when determining if a study is appropriate (or inappropriate) for use; the risks and benefits to subjects and society; and issues in determining if studies are in compliance with accepted guidelines (the complete Charge will be found in section 2.2 of this report). After generating a series of drafts, the Subcommittee met a second time in Arlington VA on November 30, 1999 to discuss various issues on which consensus had not yet been reached.

Section 3 of this report addresses each element of the Charge, and provides many specific recommendations to the EPA. The Subcommittee found, however, that its most significant findings could be best expressed outside the constraints of specific Charge issues. These findings are:

a) All research involving humans should require prior review by an Institutional Review Board.

b) The structure, function, and activities of both the Agency's IRBs and external IRBs of entities submitting data should be under active and aggressive scrutiny by EPA, with adequate staff and financial resources provided to carry out this mission. EPA should establish an internal ethics review organization to perform this function, staffed by full-time individuals whose duties address exclusively compliance oversight. The review organization should also provide an institutional focus for continuous close liaison on ethical matters with other federal agencies.

c) The Subcommittee believes that intentional administration of pesticides to human subjects testing is acceptable, subject to limitations ranging from "rigorous" to "severe." Those supporting such testing feel that the information sought must not be available via other sources (e.g., animal studies and models or study of incidental exposures), and the information expected to be gained must promise reasonable benefits to the individual or society at large. Studies should be appropriately designed to address the stated objective, and have sufficient statistical power to provide an unambiguous answer to the question under investigation.

d) In no case should developing humans (i.e., the fetus, infant, and young children) be exposed to neurotoxic chemicals. There are currently too many unknown dangers to justify such studies, even under the most extraordinary circumstances.

e) The EPA should take whatever administrative action is necessary to extend the protections of 40 CFR Part 26 to all human research activities whose results will be submitted to the Agency.

f) Some of the Subcommittee Members that accepted the use of human volunteer testing of pesticides identified certain situations in which such testing would or would not be appropriate:

1) It would not be appropriate to conduct human volunteer testing when adequate human data are already available.

2) Human volunteer studies would not be appropriate for pesticides in use today when data of equal quality can be obtained from field exposure studies.

3) Human volunteer studies could be appropriate when there are significant data gaps and such studies would provide a more accurate risk assessment.

4) Human volunteer studies could be appropriate for pesticides, which are not yet on the market, i.e. new pesticides.

5) Given the significance of statistical considerations in regard to human study design, we encourage the Agency to organize a workshop to specifically deal with this issue.

2. INTRODUCTION

2.1 Background

A Joint Subcommittee of the Science Advisory Board (SAB) and the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA) Scientific Advisory Panel (SAP) convened in a public meeting on December 10-11, 1998. The purpose of the meeting was to provide advice and comment to the Environmental Protection Agency (EPA) on issues related to data derived from testing on human subjects, particularly the use of human data for making pesticide registration decisions. Both scientific and ethical questions have been raised about the data to include the manner in which they were developed and how or whether these data should be used in the decision making process. A draft report was generated based on the presentations and discussions at this meeting. However, a significant subset of the Subcommittee remained unresolved about the content of the report. Therefore, a second meeting of the SAB/SAP Joint Subcommittee was convened on November 30, 1999 to permit further deliberations for the purpose of resolving and bringing to closure differences of opinion within the Committee.

The Office of Pesticide Programs has received a growing number of unsolicited reports of research with humans which include systemic toxicity studies to establish a human No Observable Adverse Effect Level (NOAEL). A NOAEL study is controversial in humans since dosing may be increased to a point where an adverse effect occurs. Therefore, subjects who participate in these studies will experience adverse effects with no known benefit (see section on separation of risk and benefit). Moreover, the exposure levels established by a NOAEL study may pertain only to those endpoints measured and the characteristics of the subjects who participated. Thus, the generalizability of such studies will be constrained by these factors. EPA does not require human studies to establish NOAELs and has never established guidelines for such studies in humans. As of July, 1998, EPA has not relied on the submitted human NOAEL pesticide studies to support decisions under the Food Quality Protection Act (FQPA).

The FQPA requires EPA to reassess all food residue tolerances so that by 2006 over 9,000 current pesticide residue tolerances must be reassessed. A "tolerance" is a regulation defining the allowable amount of pesticide on a food. The FQPA requires consideration of the cumulative risks of all pesticides with a common mechanism of action. This is in contrast to the previous practice of assessing exposure to one pesticide at a time. An additional tenfold safety factor must by included by EPA in risk assessments to increase protection for infants and children unless reliable data is available to support a different factor. Finally, the FQPA requires that EPA address the "worst first" pesticide. That is, pesticides regarded as the riskiest, such as the organophosphates and carbamates, are being reviewed first. Both of these classes are cholinesterase inhibitors with histories of human testing. The first third of these tolerance reassessment decisions were completed as of August, 1999 as mandated by FQPA.

Prior to registration, a pesticide must undergo many tests in animals to evaluate toxicity and extrapolate these animal study results to judge the potential toxicity for humans. These study requirements, as outlined in 40 CFR Ch 1 158-202 (e) *Hazard to humans and domestic animals*, call for data derived from a variety of acute studies and from subchronic and chronic toxicity tests. Exposure data are also required by 40 CFR Ch 1 258-202 (d) *Environmental fate* and include general studies involving fate of chosen agents, as well as studies of degradation, metabolism, mobility , dissipation, and accumulation. A reference dose (RfD) for a pesticide which is considered the "safe" daily dose is then calculated by dividing the NOAEL derived usually from the most sensitive study in the most sensitive species by a series of uncertainty factors. If, as in most cases, the study is an animal study, then a tenfold uncertainty factor is applied to accommodate variability between animals and humans. A second tenfold factor is applied to account for variability within humans, and finally the FQPA requires an additional tenfold safety factor to protect children. If, however, human data are available, the interspecies factor of ten can be dropped. Furthermore, when human data have been available and used it has generally raised the "safe dose." A higher "safe dose" allows greater use of a pesticide. Thus, the FQPA may have inadvertently created an incentive to test pesticides in humans. In fact, since passage of the FQPA, the Office of Pesticide Programs has received 14 human NOAEL studies, which represents a significant increase from the previous ten years.

For many years, EPA has performed, supported and made use of human studies on various agents of environmental concern, including pesticides, in compliance with the Common Rule. For example, EPA has required studies to determine exposure levels among pesticide applicators, mixers, and loaders of pesticides as well as field workers and others re-entering pesticide treated areas. EPA's requirements for exposure data are documented in Subdivisions U and K of its Pesticide Assessment Guidelines of 1984. However, with the increased submission of human experimental studies that involve intentional pesticide exposures, new concerns are raised regarding EPA policy for evaluating the science and ethics of these studies. Therefore, EPA convened the Joint Subcommittee of the SAB/SAP for the purpose of gathering advice to establish such a policy.

Through the establishment of "test guidelines," EPA has the authority to specify the tests required and the manner in which these tests are performed. These guidelines are established in collaboration with other regulatory agencies both in the U.S. and abroad and are subjected to rigorous peer review. EPA wants to develop a policy that applies protections, such as those in the Common Rule, consistently to all human research considered or supported by the Agency. This policy must be subjected to peer review and public comment. This policy should address the wide range of human research to include: a) incident follow-up and epidemiologic studies of humans performing usual activities; b) human experimental studies of intentional exposure such as patch tests for irritancy or sensitization, studies of pharmacodynamics or metabolism, and testing to establish a NOAEL. When the criteria for acceptability of these two classes of studies vary, EPA is requesting that the distinctions be specified. Moreover, EPA is requesting guidance in

applying contemporary scientific and ethical standards to older data or to studies from other countries.

In its initial deliberations, the Joint Subcommittee reached ready agreement on several basic and preliminary points. These include:

a) Any policy adopted by the Agency should reflect the highest standards of respect for human subjects and should prohibit research protocols that override the interests of subjects in order to obtain useful data.

b) If it can be justified at all to expose human subjects intentionally to toxic substances, the threshold of justification for such action should be very high.

c) Bad science is always unethical; research protocols that are fundamentally flawed, such as those with sample sizes inadequate to support reasonable inferences about the matter in question, are unjustifiable.

d) If the use of human subjects in pesticide testing can be justified, that justification cannot be to facilitate the interests of industry or of agriculture, but can only be to secure the public health.

e) Any policy adopted by the Agency must reflect a special concern for the interests of vulnerable populations, such as children, the elderly, and those with fragile health due to compromised respiratory function or other reasons.

f) Unintended exposures provide valuable opportunities for research; it is an error not to take full advantage of such opportunities to gain major information through careful incident follow-up.

g) In considering research protocols, it is not enough to determine a risk/benefit ratio; it is important also to consider the distribution of risks and of benefits, and to ensure that risks are not imposed on one population for the sake of benefits to be enjoyed by another. It is also important to be sensitive to the difference between a risk that may be irreversible, such as possible interference with normal neurological development.

Having agreed to these points as providing the underlying values that should inform the development of actual policy recommendations, the Joint Subcommittee then faced the challenge of providing greater operational clarity regarding the boundaries of what should and what should not be allowed. A second meeting on November 30, 1999 -- and a subsequent process of exchanging views on a developing draft of this report -- led to the conclusions and recommendations contained herein.

2.2 Charge

In pursuit of these objectives, the Joint Subcommittee was charged as follows:[1]

a) The Value of Human Studies

Human studies provide a special type of information that may contribute to the decision-making process. The Agency seeks advice on the role that such data can play in evaluating a toxicological data base for purposes of regulatory decision-making. Specifically,

1) What are the general arguments for the proper role of human studies in supplementing animal studies in making regulatory decisions about various environmental agents; e.g., water pollutants, air emissions, and pesticides?

b) Factors for Consideration

The Agency is confronted with the question of how to determine what constitutes an appropriate human study for use in environmental decision-making. There are similarities and differences between the use of such studies in reaching decisions in other areas; e.g., drug licensing. In all cases, the Agency recognizes that the scientific benefits must at least be commensurate with the risks involved.

[1] Two Members suggested revisions to the Charge to clarify the general language and to eliminate wording which could be interpreted as advocating subjecting a human subject to damage if the potential societal benefits were great enough. The Charge conveys the questions asked by the Agency, and is the starting point and framework for the Public Meeting and subsequent report. In order to maintain the historical record and process, the Charge is not changed once a meeting is completed.

1) What factors are relevant to consider when reaching a judgment on what constitutes an ethically appropriate human study?
2) How can these factors be used to make decisions in such cases? Please give some examples.
3) In using these factors, are there "benchmarks" that emerge that would clearly make a study appropriate (or inappropriate) for use? Please give some examples.

c) The Risks and Benefits to Subjects and Society

The Agency is concerned that the best scientific information be brought to bear in making its decisions. At the same time, the Agency is concerned that the studies they require/rely on to make those decisions should meet rigorous ethical standards. Specifically, the risks to the study subjects should be outweighed by the benefits for them personally or for society as a whole.

1) What are the benefits to subjects and to society from human participation in research studies; e.g., those supporting pesticide registration?

2) What is the impact of remuneration on this question of benefits to subjects and society?

3) Are there differences or distinctions that should be made for studies involving pesticides versus those involving other environmental chemicals?

d) Application to specific situations

The Agency must make judgments on a wide variety of studies involving humans. Such studies include controlled ingestion (as well as exposure by other routes) of test compounds by test subjects, accident reports, and monitoring of exposure during routine activities. It would be helpful to have advice on how the guiding principles on human subject research and testing (i.e., the Common Rule and Declaration of Helsinki) might be applied across this broad range of studies, particularly as they might apply in the case of studies submitted in support of a pesticide registration:

1) How can/should this guidance be applied to

   i) Studies conducted in the past, prior to the adoption of the Common Rule (1991), but which may (or may not) have adhered to another ethical standard of another day?

   ii) Studies gathered from the open literature for use by the Agency?

2) Is it ethical to engage in the oral dosing of human volunteers with environmental toxicants or infectious agents of interest (e.g., cryptosporidium in drinking water or organophosphates (OPs)) in order to establish a No Observed Adverse Effects Level (NOAEL)?

e) Compliance

Even if the Agency has appropriate ethical standards in place, there is the question of determining compliance with those standards. How can the Agency determine whether and to what extent its ethical standards have been met in a particular test with respect to the following aspects:

1) Informed consent
2) Voluntary participation
3) Institutional Review Board (IRB)

3. DETAILED FINDINGS

3.1 The Value of Human Studies (Issue a)

Human studies provide a special type of information that may contribute to the decision-making process. Specifically, this element of the Charge asked the Subcommittee to address and enumerate the general arguments for the proper role of human studies in supplementing animal studies in making regulatory decisions, particularly concerning various environmental agents: e.g., water pollutants, air emissions, and pesticides.

3.1.1 Information Available from Studies with Human Volunteers

Contemporary human research in toxicology proceeds from the assumption that, in most situations, we already have a considerable amount of information about the toxic properties of a given agent (derived primarily from animal research and fortuitous epidemiological studies) before we deliberately expose human subjects. However, with new pesticides (prior to registration) there are no epidemiological or exposure data available to provide a context for prediction or extrapolation. **Both researchers and regulators support the use of epidemiological and exposure data as important to the evaluation of potential environmental risks. A majority of the Subcommittee supported conducting human clinical trials with pesticides, but called for cautious approaches (i.e., that exposures must be done only under the strict ethical and safety guidelines discussed below); other Members called for severe restriction on such research, particularly when neurotoxicants were involved. One additional caveat concerning such intentional exposure is important -- the Subcommittee, in general, would not support human experimentation primarily to determine a NOAEL**. Although a No Observed Effects Level (NOEL) or NOAEL may be defined in the absence of a documented toxicological response (in which case it does not have strong scientific standing or support), such data are of value in the clinical and regulatory arena for setting exposure limits, etc. The likelihood of mechanistic insights improves with the inclusion of dosage level inducing some discernible sign of toxicity. Generating such data pose ethical concerns, however, as discussed below in section 3.1.2.

The Subcommittee believes that pharmaceutical industry practices offer useful models for human pesticides research. When a new drug is released, the manufacturer performs post-marketing surveillance, mainly to gather information about adverse effects. Similar, properly designed observational studies of humans accidentally or occupationally exposed to pesticides should be encouraged over intentional exposure studies with volunteers. These observational studies can address the nature and incidence of adverse effects in a much more diverse group than that represented by the experimental volunteers and, as such, should have greater value for risk assessment. However, such studies lose some degree of control over exposures and timing of observation that could make them very difficult to implement.

Perhaps the greatest potential value to be derived from experimental studies in volunteers is the opportunity to place the results into a structured hierarchical information base incorporating and integrating both animal experiments and human research (particularly addressing indices of neurobehavioral function in the case of insecticides) addressing short-term exposures.

Such a structured information system would provide a clearer purpose for human data. It would help overcome several of the ethical issues inherent in experiments with volunteers by providing better insights as to "safe" levels" and expected reactions. **Perhaps most crucial, within such a decision system, human experimental data would serve as a valuable transition to further research on both exposure assessment and toxic mechanisms. In such a role, human experiments would pose fewer of the ethical quandaries that arise when they are used simply to establish a NOAEL that lacks cogent scientific value and whose purpose can be interpreted as simply an argument for higher permissible exposure levels. Strategically designed studies with focused efforts and clear decision systems in place to acquire information are defensible both scientifically and ethically.**

3.1.2 Limitations of Clinical Studies

Controlled experiments with human volunteers are framed to answer a limited range of questions about the risk potential of a substance. To conform to accepted ethical standards, they are typically confined to low or moderate doses of limited duration and constructed as carefully as possible to avoid producing a serious effect, either acute or long-term. Ethical guidelines take account of both the usefulness and shortcomings of such studies, and their applicability to questions about other agents and other populations. There are several factors, discussed below, which these guidelines must take into account.

First, volunteers generally are recruited from a healthy adult population (although participation by pregnant women is not precluded by current federal policy, we believe that they should be excluded from clinical studies with pesticides, as should sensitive subpopulations such as the elderly, those with already compromised health, and children). Like the "Healthy Worker Effect" recognized by epidemiologists, such a selective process limits the generality of the findings. In addition to the "healthy worker effect," findings may be affected by the fact that some groups in society are less likely to volunteer.

Second, although volunteer experiments typically involve brief exposures, many real world questions about safety involve chronic exposures. This is particularly relevant with pesticide exposures. In one case from the insecticide literature, investigators studying a sample of farmers exposed while treating sheep with organophosphates (OPs) reported that the chronic effects of exposure, primarily neurobehavioral in character, are not predicted by sensitivity to any acute warning signs (Stevens *et al.*, 1996). Although this difference in exposure patterns can be a complicating factor and is certainly a limitation, it can often be addressed by careful experimental design, as has been demonstrated in human studies of ozone and carbon monoxide which also had to deal with the issue of brief versus chronic exposures. One Member disagreed, noting that chronic effects, such as the neurobehavioral changes seen for the OP's, would be very difficult , possibly impossible, to detect in acute studies regardless of the design.

3.1.3 Limitations on Establishing NOELs and NOAELs with Human Testing

Given the above, we must recognize that the ability of short-term human experiments to provide a scientifically meaningful NOEL or NOAEL is circumscribed, as detailed below:

Although establishment of a NOEL, NOAEL, or Lowest Observed Adverse Effects Level (LOAEL) can provide data of value in the clinical and regulatory arena, there are also ethical considerations about the research needed to establish them. The benefits of obtaining a LOAEL are discussed above. However, generating a LOAEL requires a level of exposure inducing some identifiable effect or symptom. To obtain such data raises a particular ethical problem, because it will require human volunteers to experience some toxicity-induced symptoms if the dosing levels approach critical thresholds, with no prospect of any direct therapeutic effect. This is at variance with most biomedical research where exposure to a known risk (e.g., a new chemotherapy agent) is balanced against the potential health benefits. In addition, research by Mantel and Bryan at the National Cancer Institute (1961) and later at the National Center for Toxicological Research (Gaylor, 1992) showed that even NOELs, which are statistically derived, actually correspond to some finite incidence of adverse effects. That is, for volunteers, research to identify a NOEL may not be free of risk

Testing insecticides presents unique challenges because their adverse effects are often neurobehavioral in character. If, as some reports suggest (Steenland *et al.*, 1994), such effects are more sensitive than other measures of toxicity, the use of these neurobehavioral measures might generate LOAELs at lower dose levels.

In addition, short-term volunteer experiments have yet to mimic the most common exposure pattern, consisting of repeated, intermittent, acute elevations in dose, typically to the combination of agents seen in most pesticide formulations rather than to a single agent. The degree to which intermittent or even single doses of insecticides might induce central nervous system sensitization to OP insecticides possessing proconvulsant properties is not known. Also, the scope of OP interactions with certain classes of proconvulsant medications, such as the popular selective serotonin re-uptake inhibitors, is unknown. Volunteers presumably would not be used to assay such a possibility. Whether or not conventional uncertainty factors (UF) account for the effects of such medications should be further investigated in animals and in humans exposed occupationally to insecticides.

Additional obstacles arise when attempting to extrapolate findings to children, particularly in addressing the most troublesome question in human research: the consequences of exposure during early development. Current human volunteer studies are not designed to yield a reference dose for children, but rather (as noted before) to place some portion of the animal data into human context. The biology of the child diverges markedly from that of the adult. This difference is probably best seen in the central nervous system. Before, and for a number of years after birth, the child's nervous system develops at an extremely rapid rate. Nerve cells are laid down and migrate to their final destination; connections are built; synapses are formed; and neuron populations are pruned. Perturbations to the nervous system at this time may produce persistent changes in brain architecture. The particular sensitivity of the developing organism to insult has been shown for so many noxious agents that it has achieved the status of a general principle. Among the exemplars are the effects of oxygen on the premature infant eye and the effects of prostaglandin antagonists on closure of the ductus arteriosu. These, and the effects of lead, mercury, alcohol, dilantin, bilirubin, and cocaine on the infant brain establish this principle **(Need Ref here)**. Dosing healthy adults provides extremely limited (if any) insight into the risks for the developing brain.

Such a distortion of the response profiles may not be fully accounted for by the imposition of traditional UF when results are extrapolated to the general population. The conventional UF of 10 for inter-individual variation dates from the 1950s, and is not an instrument devised to reflect contemporary molecular toxicology.

The magnitude of an intraspecies UF based on rodents also has limited bearing on the appropriate UF for children. Furthermore, neurotoxic insecticides induce many effects on the body and nervous system. Each is characterized by its own dose-response function. Attempts to establish a NOAEL on the basis of a single outcome, such as peripheral acetyl cholinesterase (AChE) levels, may mask a substantial Type II error. That is a type of error causing the null hypothesis to be improperly accepted, so that an effect which is actually present is not identified (e.g., a neurotoxic effect of an OP that occurs at a lower dose level than would cause a statistically significant change in a measure such as peripheral AChE levels). For example, although cholinesterase inhibition by carbamates is rapidly reversible, the symptoms of toxicity may linger, so that cholinesterase assays in this instance may provide an erroneous diagnosis.

Further deterrents to extrapolation from volunteer studies to children are posed by two additional factors that have led EPA to conduct targeted exposure assays:

a) Young children occupy a different spatial ecology than adults. They often experience elevated exposures simply because their environment lies close to floor level. With metallic mercury, for example, vapor concentrations at floor level may be 10-20 times higher than concentrations at waist level. Dust stirred up by activities such as crawling causes increased inhalation of lead dust and pesticides residues by children.

b) As a result of the spatial niche they occupy, young children have a propensity (as a function of their close proximity to the floor or ground and/or behavior in crawling) to either pick up or be exposed to objects or substances on the ground. They tend to explore their world by hand-to-mouth sampling, which increases their exposures considerably. Such behaviors help explain why children living adjacent to agricultural sites tend to experience elevated pesticide exposures. Adult NOAELs, obtained under highly controlled conditions, have to be modified to account for such exposure sources. This problem is recognized by the FQPA in its requirement to aggregate total exposure from all sources, which may diminish the usefulness of volunteer data.

In any study involving potential harm to the study participants, whether humans or animals, there is an ethical necessity to be certain that the study has sufficient statistical power and is appropriately designed to address the objective of the study. Many Institutional Review Boards (IRBs), in fact, now require documentation that the proposed study possesses adequate statistical power. This is a multifaceted issue requiring consideration of a number of factors, which are detailed in Appendix A.

The most serious problem of those identified above is that of generating data applicable to the developing child (or fetus). There seems little probability that high quality data relevant to children can be derived from studies on adults at this time, or in the foreseeable future. The Subcommittee rules out the only alternative, the testing of children, as being ethically unacceptable. There are too many unknown dangers to justify the effort, even under the most extraordinary circumstances.

Despite the constraints, uncertainties, and risks noted above, experiments with human volunteers can still provide helpful information. With radioactive isotopes, they can help trace the distribution pattern of a chemical and its persistence in certain organs, as with mercury. They can help determine if specific subpopulations are predisposed to adverse effects from acute exposures, as with the response of asthmatics to air pollutants. They can help determine the relationship between exposures and exposure biomarkers, as with the correlation between specified doses of organophosphate insecticides and cholinesterase levels in blood. Volunteer experiments with pesticides can be useful as guides to additional laboratory research with animals and the formulation of more specific animal models.

3.2 Factors for Consideration in Identifying Ethically Appropriate Human Studies (Issue b)

The original Charge posed three specific questions:

a) What factors are relevant to consider when reaching a judgment on what constitutes an ethically appropriate study?

b) How can these factors be used to make decisions in such cases?

c) In using these factors, are there benchmarks that emerge that would clearly make a study appropriate (or inappropriate) for use?

Because these questions are closely intertwined, the Subcommittee has chosen to address them collectively, focusing on the following factors.

Study Design: The Subcommittee unanimously supports the principle that any study that does not have a clearly defined hypothesis and proper study design to test that hypothesis is *per se* unethical.

The EPA relies on the determination of a no-adverse-effects level (NOAEL) and/or a lowest-observed-adverse-effects level (LOAEL) in setting reference doses for toxicants. This procedure raises serious concerns about the ethical use of human data in the evaluation of health risks of environmental hazards. An experiment that does not have a chance of achieving its goal, in this case estimating the effect it seeks, is *per se* unethical.

Considering the other problems associated with the use of NOAEL/LOAEL's (e.g. design dependency, not an estimated value but the result of a test), the Subcommittee does not believe human studies should be used to directly estimate these quantities. However, a properly designed human study with sufficient sample sizes could aid in understanding differences in metabolism and help to guide the species extrapolation. Appendix B presents a detailed discussion of how sample size affects the ability of a study to detect small changes and effects.

The Subcommittee believes that issues of age, gender, and ethnicity should receive consideration in designing studies and assessing their relevance for regulatory purposes. Though the Subcommittee opposes the use of children as experimental subjects particularly in relation to intentional exposure of children to toxic agents, it also supports the concept that the relevance of studies to assessing the risk to children should be specifically addressed. Special concerns were expressed that risks to developing organ systems might be less reversible than to mature systems and that the risk to children is unacceptable. This concern also would affect the potential ability to generalize from adult subjects to children. Likewise, studies performed in male subjects must be examined to determine their relevance to female subjects. Ethnic variation in response must also be considered.

Overall Considerations: Existing federal standards, noted below, can serve as an initial guide to discussing the Charge questions. These standards, however, basically apply to drug development protocols. In this model, research is guided by the premise that its eventual goal is either to benefit the subject directly or to benefit patients with a specific disease. Because it presumes the possibility of benefit, it weighs the risk of possible harm against potential benefit. The ultimate aims of volunteer testing for drug development and for pesticide exposure standards diverge because of the benefit component of the risk-benefit equation. **That is, both drug and pesticides testing have financial goals and both use healthy volunteers who do not stand to benefit personally. Thus, the risk and benefit are split. They diverge in that drug studies can be easily justified because they benefit others with a disease or condition which cannot be said for pesticides. However, protection of the food supply has a societal benefit that we do not see for drugs. In the course of marketing, drugs are targeted to a specific population in need and their effects are monitored by physicians. However, in the case of pesticides, a broader population is potentially exposed and not monitored for health effects. This situation is a powerful argument for the conduct of controlled exposure studies to better understand the effects of low level exposures. Otherwise, the populace is our controlled exposure study.**

The intention variable: A core question for ethical review of a proposed or submitted study is intent. Although intent might be argued as beyond the purview of an ethical review, and is difficult to interpret, the Subcommittee views it as a critical issue, in part because it helps define the scientific value of a study. For that reason, it maintains that the intent of a proposed study should be defined clearly at the outset. It agreed that, generally, human dosing experiments are not appropriate if the primary intent of the study is to determine or revise a NOEL or NOAEL so as to eliminate the interspecies uncertainty factor. Studies designed to advance scientific understanding, for example, to clarify mechanistic questions, may be ethically defensible. A cogent model for such experiments would be the studies of mercury vapor conducted by Cherian *et al* (1978). These investigators had subjects inhale trace amounts of ^{203}Hg, then followed the time course of its distribution in various tissues and fluids. The experiment was not designed to provide any direct information about toxicity. The use of observational or epidemiological studies to test hypotheses that are appropriately addressed by such studies often present only very limited ethical concern. The test of intent for ethical "acceptability" resides in the scientific value of a study, i.e., its potential to provide useful information. Although rigid rules should not be imposed, a weight-of-evidence approach should prove useful. For instance, as discussed elsewhere in this report, does the study have sufficient statistical power? Is more than one dose included? Are sensitive and comprehensive response indices described? Do they extend beyond conventional clinical observations? Could the results be extrapolated to the population at large, considering its age, genetic, gender, and ethnic diversity? The enormous breadth of such variation, which contrasts sharply with the typical volunteer pool, presents difficult problems for extrapolation, as discussed elsewhere in the Subcommittee report. Would the research be acceptable in a recognized peer-reviewed journal?

Standards for risk review: The discussion below amplifies the Subcommittee's application of existing Federal guidelines for human subjects research. It might be framed as a set of questions to an IRB.

a) Have risks to subjects been minimized?

IRB reviews must devote considerable care and attention to items such as medical exams (including histories) to determine the health status of the subjects, identifying medications taken by the subject and alcohol use. The review should also examine the doses or concentrations to which subjects will be exposed and determine how these relate to our existing knowledge of the agent's effects. It must also examine the plans for dealing with any unexpected response to the agent administered. The IRB needs concrete details, and should assure itself that these have been provided in sufficient depth.

b) Are risks to subjects reasonable in relation to anticipated benefits?

1) Is there an important research question being asked, one which could not be addressed with animal data?
2) Are the design instruments and methods, and the competence of the researcher, appropriate to answer the question?
3) How will the research findings be used?
4) Have designs which could pose less serious ethical issues been considered (and if rejected, why)?
5) Is there a need to use human subjects?
6) Have stopping rules been described?

Noting other issues that arise when one considers the scientific merits of an experiment may help to illustrate further the issues raised by the Charge. There may be a difference between the scientific assessment (today) of a study, involving the evaluation and use of existing data (e.g., a retrospective review of previous exposures, where consent had not been obtained), and prospective studies proposing intentional exposure of human subjects to a test agent. How should IRBs judge current studies using different and more rigorous standards than applied to the acceptance of studies conducted in the past under somewhat different ethical guidelines. How should they evaluate retrospective studies, consistent with 40 CFR 26.119?

Should IRBs require of those who use historical data to specify the data's origins, methods, and limitations, especially where concerns about the validity of those studies exist?

c) Is the selection of human subjects necessary and appropriate?

Does the question asked require human testing? Before undertaking human experiments, one should carefully decide whether the information one plans to obtain can be derived from animal studies. This is particularly true with "new" unregistered pesticides because human exposure information will not be available, and estimates of risks to humans will have to be calculated on the basis of animal studies. Even if the estimates from animals are highly uncertain, calculations of these values are critical to the proper design of the human investigation. These provide target estimates for the measures of interest and valuable guidance as to what might be expected in the human studies.

d) Have less ethically questionable studies been considered?

Controlled exposure studies in humans are problematic in that they raise ethical questions. In some cases, epidemiological studies or studies of exposed populations may be able to obtain/provide virtually the same information, or at only a modest cost in the relevance of the information. However, for "new" pesticides, epidemiological information will not be available.

e) Is the informed consent process properly designed, with the opportunity for potential subjects to think through the relevant issues, including possible compensation for harm?

Related issues: Some aspects of risk assessment that bear on the ethics of human testing were not addressed by the charge nor are they properly addressed by prevailing Federal guidelines. They are pertinent to the Subcommittee's task, however.

a) Ethical questions pertaining to human testing range beyond conventional risk assessment evaluations. It would be advisable to include contemporary risk characterization issues such as individual and community risk perceptions and acceptability (Stern and

Fineberg, 1996). The role of the community in research involving human subjects is now generating considerable interest in the clinical trials world. In response to community concerns, potential subjects may be invited to participate in the design of studies; or, community input may be sought, directly or indirectly, about the use of research data. Also, community concerns about potential benefit and harm may be surveyed. EPA , in fact, is increasingly attentive to this broader context of risk characterization. This is true with genetic studies to determine susceptibility, but it is also true as we begin to explore the connections between genetic and environmental factors in the etiology of human disease.

The Subcommittee recognizes that such an expansion of ethical dimensions may present difficulties arising from risk perceptions in a particular community. For example, the public historically has been more concerned about cancer than other potential risks, and judges its adverse effects to exceed those associated with other potential risks. Both investigators and IRB members need to be sensitive to public perceptions, however, to acknowledge them in informed consent documents, and to guard against the intrusion of their own values and perceptions into their evaluations.

b) The Common Rule has a specified and very helpful list of required considerations concerning informed consent. Because of questions raised by past studies on pesticides, the Subcommittee notes that a useful way of determining whether a potential subject grasps the information in the consent form is to administer a brief multiple-choice test based on the form's content. Such a test provides a measure of how well the subject grasps the contents of the consent form.

c) Other issues that may be of particular relevance to both environmental agents and pesticide testing include:

1) Although subjects should have the undisputed right to withdraw from a study at any time, exercise of this right could make it difficult for researchers conducting environmental exposure studies, particularly longitudinal, epidemiological studies, to complete their protocols. This right must, however, be described unambiguously.

2) Subjects should have rights to compensation if they are injured as a result of the experiment. Since injury may only become evident long after the experiment,

such compensation issues need to be addressed at the inception of the study. As part of their reviews of experimental protocols, IRBs should request the investigator's plans for ascertaining the subjects' health status for some period after the end of the experiment, and ensure that each subject is given clear information about how to deal with problems that might emerge later. .

3) General issues related to privacy rights and confidentiality are already described in existing regulations. Additionally, there are specific concerns about the use of confidential information obtained from a subject's participation in a study. For example, the use of data relating to susceptibility to certain diseases that have an environmental component (e.g., paroxonase levels) may place individuals at risk of discrimination (health care, life insurance, employment). These issues would need to be addressed in the consent process and protections built into the protocol.

4) University-based research has been displaced in many instances by contract organizations dependent on relationships with industry clients. These relationships may arouse skepticism about the assumption that the experimenters are neutral parties. Moreover, such relationships also provoke concerns about the IRBs appointed to review study protocols. These include the criteria for membership on an IRB (inclusion of public members, advocacy groups, etc.); criteria for approval (consensus vs. voting); and public disclosure of reasons for decisions.

5) When the results of volunteer studies are submitted for publication in scientific journals, it is essential that the sources of research support be disclosed unambiguously. Several prominent medical journals have encountered possibly deceptive statements about such support.

6) Independent review is especially crucial, but an increasing number of private IRBs are now operated by commercial, "for profit" entities -- an environment that may pose problems when attempting to conduct a truly independent review, and that calls for close scrutiny. At the same time, it should be recognized that the number of privately operated IRBs has increased because of the financial and operational efficiencies they offer. In addition many academic institutions may lack the resources to conduct the appropriate

reviews and to fulfill administrative requirements imposed by the federal government and other oversight authorities. Monetary compensation for members of private IRBs, however, should be described in any submission to EPA.

In view of the complexity and interrelationships of the manifold questions presented by volunteer studies, the Subcommittee agreed that no specific benchmark, algorithm, or unambiguous dividing line could be applied universally to categorize research as either unethical or ethical. Such judgments require the weighing of multiple factors in two categories: technical and scientific issues (e.g., sample size, experimental design, and the nature of the agent under study), and subject welfare issues (e.g., provision of informed consent, lack of coercion of any type, and compensation for any harm done in the course of the experiment).

Because of the lack of fixed landmarks, except perhaps at the extremes, the Subcommittee proposes that the Agency offer guidance in the form of examples. One extreme at the innocuous end of the scale might be exemplified by a skin irritation study with glyphosphate in adult males. The other extreme might be exemplified by a study designed to obtain a NOAEL for neurotoxicity with a highly potent organophosphate. The territory between these extremes is where the Agency needs to provide guidance both for its own policies and for parties contemplating the submission of human volunteer data. Appendix C is one suggested model. The scenarios there could serve as the basis for asking questions such as:

a) Who would be acceptable volunteers? Under what conditions, if any, are the aged, and female subjects acceptable?

b) What is the hypothesis? What would be the intent of the study, e.g., kinetics, determining LOAELs, etc? Under what conditions are these studies appropriate/acceptable?

c) Given the intent, how would reasonable sample sizes be determined?

d) What level of dosing is appropriate, acceptable? Are there conditions under which dosing to measurable/observed toxicity is appropriate?

e) If dosing can be administered to toxicity, what organ system toxicities are acceptable? Can neurologic toxicity ever be accepted? Can or should biochemical alterations be used as surrogates for clinical toxicity? Do these surrogates allow/promote protection of subjects?

3.3 Risks and Benefits to Subjects and Society (Issue c)

Issue (c) of the Charge posed three separate, but interrelated, questions concerning both the risks and benefits associated with human experimentation. Each of these questions, and the Subcommittee's responses, are addressed below.

3.3.1 The Interrelationship Between Science and Ethics and the Benefits of Research Involving Human Subjects

Interrelationship between science and ethics: The design and conduct of research involving human subjects involves two types of considerations: First, research must have scientific merit-it must ask important and relevant research questions that have not already been adequately answered, and must do so based on a rigorous methodology that can answer the research questions. Second, research must be ethically acceptable-it must be based on a set of ethical considerations that provide assurance that the rights and interests of subjects will be protected and that valuable and important research will be conducted.

It should follow that there is an interrelationship between science and ethics -- a research design that does not deal with a novel, important and relevant question, or is not based on rigorous scientific methodology (or both) cannot be considered good from either a scientific or an ethical perspective. Indeed it has been said that good science is a prerequisite for good ethics. This is more than just a statement of intent or of aspiration. The separation of science and ethics-as occurs when scientific peer review precedes the evaluation of a study by an Institutional Review Board (IRB)-may be procedurally necessary, but it is a separation that is arbitrary and difficult to defend.

For almost all scientific considerations in the design and conduct of a study there are ethical counterparts, and vice versa. For example, the scientific requirement that a study is well defined, asks novel questions, or can obtain measurable outcomes can also be seen through an ethics lens: is the study necessary? Is the research question important? Are the needs of potential subjects and/or society being met? Both sets of considerations relate to the importance of the research question. Similarly, the questions one would ask from the ethical perspective (Is the study feasible as designed? Has there been fairness in the recruitment/retention of subjects? Are appropriate safety procedures in place to minimize potential harm to subjects?) are relevant to the scientific requirements that a study be well designed, that the study architecture is appropriate (e.g., a case-control study, and randomized trial), that methods have been introduced to reduce bias by investigators, and that the methods of monitoring procedures in the case of toxicity, drop out by subjects or discontinuation are appropriate.

Attention to the interrelationship between science and ethics in research involving human subjects need not involve simultaneously assessing both; rather, it requires only an appreciation that scientific merit and ethical acceptability are jointly necessary conditions to be satisfied prior to enrolling human beings in research. Investigators, IRBs, and regulatory agencies should not lose sight of this crucial principle.

The benefits of research involving human subjects, e.g. those supporting pesticide registration: The question of benefits to human subjects needs to be considered as part of the assessment of risk in relation to potential benefits (sometimes referred to as the "risk-benefit ratio) posed by human subject research. Human research subject protection policies evolved out of experiences in which research subjects were exploited, such as in the Tuskegee Syphilis Study. In addition to the outright deception it involved, the Syphilis Study further exploited research subjects in that all the burdens of research were borne by the subjects, and all the (future) benefits of the research were realized by others. This splitting apart of risks and benefits is not necessarily unethical, but raises the potential for exploitation and thus argues that ethically acceptable levels of risk ought to be lower when there are no offsetting potential benefits to subjects.

Research in which risk and benefit are separated in this way was historically known as "non-therapeutic," and is now referred to as 'research without prospect of direct medical benefits to the subject.' It is therefore important to enumerate and understand the potential benefits to subjects of research supporting pesticide registration, since identification of potential benefits to the subject, if any, will influence assessments of the ethically acceptable levels of risk imposed by such research.

This section will examine the risks and potential benefits to be realized by subjects in research that supports pesticide registration. Given that human dosing studies for pesticides expose research subjects to risk in an effort to gain information in support of pesticide

registration, it is important to identify whatever potential benefits might accrue to subjects as part of making an appropriate risk-benefit assessment.

a) **Research with potential medical/health benefits to the subject**: Some research in support of pesticide registration may have direct medical benefits to research subjects. This includes studies in which the subjects have been exposed to pesticides for purposes other than research, such as in occupational settings, and accidental or environmental exposures. The potential benefits may accrue to both the individual subjects, their families, co-workers and the groups and communities to which they belong. Research subjects themselves may benefit through increased health monitoring, safer work environment, and improved protections (protective clothing, respirators, etc.). It is, however, important to distinguish between direct benefits (those benefits that are the direct result of the intervention itself) and indirect benefits (those that arise as a later consequence of the intervention. As noted below, many of the potential benefits from pesticide testing for purposes of registration may be of the latter type.

In addition to whatever benefits may be realized by subjects themselves, there are potential benefits to those with close relationships to the subjects that must also be taken into account. The family of research subjects may benefit in cases where family members live near the workplace and are exposed to similar hazards as the individual (e.g., farm workers). The benefits to family members may include health monitoring, safer living conditions, and improved protections.

b) **Research with potential direct medical or health benefits to groups to which subjects belong, but not to the subjects themselves**: Some research offers no prospect of direct medical or health benefits to the subjects themselves, but may benefit groups to which the subjects belong. Potential benefits may accrue to groups that may be at increased susceptibility due to genetic variation, e.g. an ethnic group to which the subject belongs, but for whom the research will not have a personal impact. The benefits of research may accrue to future workers in similar occupational settings to the subject, e.g. fellow pesticide workers/loaders, even if the research will not have a direct impact on the research subject. And the benefits of research may accrue to the members of a geographical community that has been exposed, and whose members will benefit from a safer environment in the future, even if the research will not have a direct impact on the research subjects themselves.

In the case of testing of pesticides, particularly those used in human food production, there is potential for the volunteer to benefit, as a member of the general population, from participating in such a study because of that person's potential for being exposed to the pesticide in his/her food. However, in the

case of pesticides, a broader population is potentially exposed and not monitored for health effects. As noted earlier, this situation is a powerful argument for the conduct of controlled exposure studies to better understand the effects of low level exposures. Otherwise, the populace is our controlled exposure study.

c) **Research with potential medical or health benefit only to the population at large**: Some research offers benefits to subjects in only the most removed sense, through benefits to the population at large. The general population may realize benefit from increased safety information and a safer environment due to the information yielded by human testing of toxicants. But research that yields benefits to the population at the expense of risk to the subjects of research is ripe for exploitation, and may arguably be inherently exploitative. Moreover, unlike the potential benefits described above, benefits to the population at large may only accrue at a future time.

Further, the economic benefits of pesticide registration should not be considered in the risk-benefit ratio of pesticide testing on humans, any more than the economic benefits of pharmaceutical development ought to be considered in the risk-benefit ratio of new drug testing. Lastly, payment or other remuneration for participation in research should not be considered as a benefit of research to be weighed in offsetting the risk posed. As discussed below, the level of remuneration should never be so great as to encourage overlooking the risk imposed by the research, or to compensate for it.

Potential benefits are, like risk, often difficult to predict with accuracy, especially for individuals.

3.3.2 The Impact of Remuneration on Benefits to Subjects and Society

Element (c)(2) asked the Committee to identify and discuss the impact of remuneration on this question of benefits to subjects and society.

Remuneration for volunteer studies arouses debate because it can be a form of undue inducement. IRBs confront this issue repeatedly, even when volunteers face minimal risks. Many are reluctant to permit large sums to be offered under these circumstances. There is no fixed standard, however, and prescribing such a standard is not feasible, given the varying situations in which subjects are recruited. For IRBs, a review of the remuneration to be offered to subjects is partly subjective, and partly governed by community standards. Similarly, the degree to which remuneration becomes unethical will depend upon both community and individual criteria.

Figure 1 describes community standards with such a model. Three counties, representing three different median family incomes, were selected from U.S. Census data (Bureau of the Census, 1993). The figure depicts different standards that might be used by a local IRB. In the poorest county, $40 per

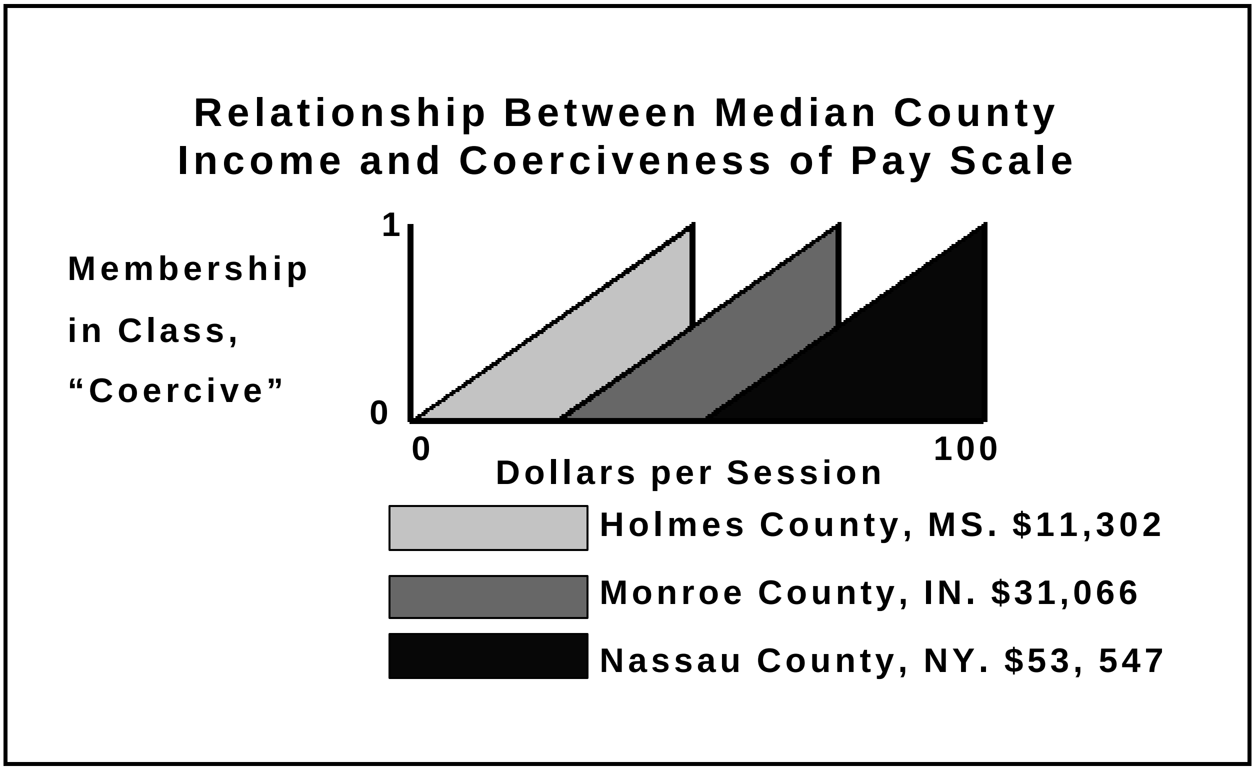

Figure 1. Depiction of how community economic standards might determine the degree to which fees for participation in experiments represent coercion.

experimental session might be judged as excessive and potentially exploitive. For the wealthiest county, potential exploitation might occur only at a $100 fee. It would be up to the local IRB to make the judgment, and various decision analysis techniques could be applied to calculate non-exploitive levels of remuneration representing a broad consensus of the IRB.

3.3.3 Distinctions Between Pesticides and Other Environmental Agents

The third element of issue (c) asked the Subcommittee to consider if differences or distinctions should be made for studies involving pesticides versus those involving other environmental chemicals.

This discussion focuses on the use of intentional, controlled, human exposures to gather data (e.g., pharmacokinetic information) on the agent(s) under study. Pesticides do not stand alone as environmental chemicals that have been intentionally administered to humans to determine the dose at which health effects occur. For example, an extensive literature documents controlled inhalation studies in which humans have been exposed to organic solvents from minutes to hours, under sedentary or exercise conditions, at varying doses during which uptake, metabolism, subjective symptoms, physical symptoms, neurologic signs, and behavioral performance have been measured **(Ref. to be provided)**.

It is important to understand what a pesticide is and how it compares with other environmental chemicals. Chemically, there is nothing unique about a pesticide; what makes a chemical/compound a pesticide is its use. Pesticides do not all share the same chemistry, toxicity, use, mode of action, or measurable health effect. Therefore, it is not accurate to discuss pesticides as one class of chemicals. In addition, the same chemical can be a pesticide in one case, an "inert" ingredient for a different pesticide in another case, and even a food additive in other cases. It should be noted however, that pesticides are, as a class, applied in many ways, including spray planes and fogs. Pesticides not only cover the target area, but, to a large degree, drift from it, exposing unsuspecting individuals and wildlife. Therefore, pesticides are, in this sense, unique. Both target populations and non-targeted populations may receive a dose. These agents, liberated around homes, cities, agricultural fields, etc., have unique relevance when both target and non-target systems are involved and often impacted.

As for any study, the risks must be weighed against the possible benefits. If it was relatively clear that no specific benefits would accrue for the individual subject exposed to pesticides in a controlled experiment, the motivation to participate in such a study could arise purely from the desire to benefit humankind (although in a coercive environment the motivation could be of many sorts apart from altruism, especially in a context of implicit coercion (e.g., the desire to avoid reprisals for being uncooperative in the context of a subject's employment)). **In fact, as noted at several points in this report, a person volunteering for human studies involving pesticides used on food has a greater chance of personal benefit from such research than a person participating in Phase I clinical studies of a pharmaceutical. We all have the potential to be exposed to a pesticide in our food (including the volunteer) versus a person in Phase I testing, where only one of ten**

drugs at that point of the review process ever reaches the pharmacy shelf.

Is it possible to conceptualize what can be learned from controlled exposures to pesticides as a benefit to humankind? If such studies were part of a program of research in which the controlled human exposure was built on extensive animal data and the purpose was to administer the lowest dose possible to humans for perhaps the purpose of validating a subtle neurobehavioral health effect, then a benefit could be construed. Moreover, if the detection of such a health effect led to reduced use of pesticides, then the benefit of less pesticide in the environment could be realized. While food producers may not regard this as a benefit, it seems likely that society as a whole, given the concern for an environment and a food supply with fewer environmental pollutants and chemicals, would construe this as a benefit. However, if the purpose is primarily to support the monetary gain of a company marketing a product with no ability to rationalize the exposure in terms of general benefits to society, then the risk to individuals does not support this benefit.

Based on these considerations, the overall conclusion appears to be that there are no specific toxicological grounds on which to differentiate pesticides from other environmental chemicals. However, pesticides may be differentiated from other agents in that the whole population is potentially exposed through ingestion of residues in food and many inadvertent sources such as those arising from spraying, deposition in household dust, and drinking water. **In addition, we typically attempt to limit exposures to all other environmental chemicals, whereas with pesticides there is a constant tension between the desire to obtain enhanced toxic effect on target pests (and possibly to increase profits to the manufacturer), and efforts to limit exposure to non-target organisms. The major motivation for such testing is usually to bring a product to market or to address a specific regulation. Thus, we are left with weighing the risks to individuals of a particular exposure against the benefits to society and the environment as a whole to decide whether an individual controlled exposure is ethically justified. Some Subcommittee Members contend that, unlike drug testing, no personal benefit can accrue to a subject intentionally exposed to pesticides or other environmental toxicants; others disagree , noting that there can be benefits, e.g., reduction in future exposures.**

3.4 Application to Specific Situations (Issue d)

3.4.1 Judgment of Current and Past Studies

The first element of Charge issue (d)(1) asked the Subcommittee to advise the Agency on how to apply ethics guidance to a) human studies conducted in the past, prior to the adoption of the

Common Rule (1991) but which may (or may not) have adhered to the ethical standard of the time, and b) to studies gathered from the open literature for use by the Agency.

For the Environmental Protection Agency, the concept of the "ethical standard of the time" is anchored by language in Public Law 92-516, the October 21, 1972 amendments to the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA). The 1972 statute makes it unlawful for any person to use any pesticide in tests on human beings unless such human beings a) are fully informed of the nature and purposes of the test and of any physical and mental health consequences which are reasonably foreseeable therefrom, and b) freely volunteer to participate in the test. The impact of this statute is to specifically prohibit, from 1972 on, use by the Agency of test data derived from human studies, unless any human test subject voluntarily exercised his or her informed choice. On the other hand, the 1972 statute must also be viewed as permitting use by the Agency of test data derived from human studies, when the law's strictures are met.

With regard to data derived prior to enactment of Public Law 92-516, the Subcommittee agreed that the fact that research was done unethically does not alone require rejection of the results of that research.

Some useful starting presumptions as we consider this issue of possibly unethical research are:

a) Useful data may, and often should, be used when they have been collected in compliance with any applicable law or regulation.

b) We ought assiduously to protect the public health and the environment.

c) We must condemn unethical research and seek to prevent it.

d) Poorly designed and/or executed research is unethical science, regardless of other “traditional” ethical considerations.

In addition, it is at best imprudent to ignore the data yielded by accidents and catastrophes. We have learned much from mass chemical disasters, e.g., the epidemic of methyl mercury poisoning in Iraq in the winter of 1971-72, which led to the discovery of the sensitivity of the fetal brain to this chemical. This incident spurred developmental research. Considerable information has also been gained as a result of studying the survivors of the World War Two atomic bombings, and persons exposed to radiation by the Chernobyl disaster.

"Incident reports," accidents, and unanticipated problems involving risks to people should be documented rigorously, and victims should be monitored afterward for some time. This documentation and monitoring is especially important where children are involved.. Such monitoring should be the responsibility of the manufacturer of the toxicant, but it is unrealistic to expect that this responsibility will be consistently met on a voluntary basis, so some manner of government requirement is essential. Also, in comparing deliberate with adventitious exposures, note that the Common Rule implies that the former prescribe a much greater degree of scientific rigor. With adventitious exposures, we are quite aware of the flaws in the data. With experimental data, the flaws may be subtle and not conveyed directly in reports, so that EPA has to be more alert to them.

When we consider research results that have been obtained in a manner inconsistent with accepted ethical standards, it is important to ask several questions, including:

a) How serious was the ethical violation? There are varying degrees of ethical deficiency. Research that conscientiously adhered to then acceptable ethical standards might not be acceptable today under more stringent, current standards, but it is not equivalent in its violation of ethical principles to research that callously disregarded the ethical standards of its time (Caplan, 1992, 1993, and 1998, discusses this topic in detail).

b) What is at stake? Is the use of the results of substantial benefit to the public health, or is the benefit simply commercial? If the intended benefit of the use is protective, it is far easier to justify that use than if the intended use is to support the release or approval of a product.

c) Are there alternative sources of or routes to equivalent information? If ethical animal studies or other human subjects research can serve the same purposes, then there is no need to rely on ethically tainted data.

If the answers support using the results, additional considerations come into play. The users of the research should issue a strong statement, explaining why the use is justified and unambiguously condemning the ethical violations associated with the research. In this way, the use of the data can be made into an opportunity to teach and to reinforce the ethical standards that must be observed by future research.

When subjects of the research are still accessible, it is best to consult with them about the intended use of the research. This may open the door to compensation issues, as in the case of the human radiation inquiry (ACHRE, 1995), but is also an important affirmation of the respect for human subjects that is at the core of ethical research.

In addition to considering what is to be done with the data and compensating (if possible) the subjects via remuneration and/or medical-follow up (the need for which may not have been known prior to discovery of the unethical research), we must ask also what to do about the fact that the research was done. It might be appropriate, or necessary, to identify and sanction the researchers (with criminal, civil, or professional penalties -- fines, barriers to or bans on future funding), the institutions where the research was done, or the financial supporters of the research. And of course, the discovery of unethical research can always raise questions of a need for revision of policy.

Some journals demand convincing evidence that research results submitted to them have been obtained in ethically appropriate ways. When they reject manuscripts for inadequate documentation that ethical standards have been met, they play an educative role that helps sustain the integrity of research. But this cannot be assumed of all journals, nor even of all that have respectable standards of scientific quality. So each study must be evaluated on its own merits before a conclusion about its ethical propriety can be warranted.

No algorithm can exist for making the decisions raised by this question. One can draw a temporal "bright line" benchmark, affirming that from a certain date, all research must meet certain ethical standards to be accepted by the Agency -- no matter who has done it, where it was done, or how it was financed. But for prior research, as well as incidents (e.g., the methyl mercury poisoning in Iraq), there is an unavoidable need to rely on judgment. For this reason, it is crucial that there be an on-going capacity in the Agency both for providing supportive advice and guidance to researchers and for scrutiny and oversight of research activities.

There is a further value in elevating the ethical standards of research: unethical research is often methodologically flawed as well, sometimes in subtle ways. For example, if the executives of a corporation, or in a research community, are invited to volunteer to participate as research subjects, they may agree because of subtle contextual coercion -- a sense that they are disloyal to their employer if they decline, or that they will forgo good favor that may matter to their future. Recruiting them as subjects is therefore ethically objectionable. For parallel reasons, they may be less likely to report adverse outcomes than subjects who have no other connection with the research enterprise, and the results of such research are thus methodologically tainted as well as ethically flawed.

The point of raising such concerns is not to eliminate research or even to impede it unduly, but to prevent the abuse that occurs when subjects (even if they are not harmed physically) are induced to participate in research in the face of risks they do not properly understand.

There will, of course, be transitional issues even if the Agency takes an unambiguously clear position for the future. Some studies may already be well underway that fail, perhaps narrowly, to satisfy strict ethical standards. And it will take time, effort, and investment to convey to all relevant constituencies just what it takes to conduct research with sufficient ethical sensitivity to meet the highest standards.

3.4.2 Oral Dosing

The second element of Charge (d) asked the Subcommittee to comment on the ethical issues attendant to the oral dosing of human volunteers with environmental toxicants or infectious agents found in the environment (e.g., cryptosporidium in drinking water, or organophosphates (OPs)) in order to establish a NOAEL. Since the Agency must make judgments on a wide variety of studies involving humans, it would be helpful to have advice on how the guiding principles on human subject research and testing (i.e., the Common Rule and Declaration of Helsinki) might be applied to a given study particularly as they might apply in the case of studies submitted in support of a pesticide registration.

Comparing oral dosing as a route of human exposure for environmental toxicants with other routes, it seems apparent that, from a toxicological standpoint, it is inappropriate to consider oral dosing any differently from the other two possible routes of human exposure to pesticides, e.g., inhalation or dermal exposure. It is clearly pointed out in the Agency's guidelines that, when testing xenobiotics in animals, the route that most closely mimics the route of human exposure of concern should be used. In that regard it would normally be appropriate to use inhalation as the route for estimating the hazard to an applicator of the pesticide or person downwind from a spraying operation. Similarly, it would be more appropriate to use dermal exposure for the same pesticide if one is interested in the hazard from working in a field at some point after that same spray operation. Following this example to its logical conclusion, the most appropriate route of exposure for ascertaining the toxic potential of that same pesticide as a residue on food would be to use an oral exposure.

One could appropriately design a study to evaluate the absorption, distribution, metabolism, excretion and pharmacokinetic behavior of a given chemical in humans with some assurance that exposures were below the NOAEL. Such an exposure (one would not need to use multiple doses) would automatically become the NOEL in this context. Obviously, one would not know if the NOEL was potentially higher but one could say with certainty that a given dose, by definition of the term, was a NOEL.

The Subcommittee's discussion at the public meeting centered on pesticides, and did not address infectious pathogens, (e.g., Crytosporidia, as called out in the Charge). It was recognized during the development of this report, however, that studies of infectious pathogens also must be carefully considered in terms of their potential hazard to the volunteer as contrasted to the potential benefit to society at large. Such a study would require a very high justification. The basic difference between this type of study and one of a chemical nature in terms of "dose" is that a study of an infectious agent provides only two endpoints: infection (disease) or no infection (no disease). If the former result is encountered, that particular individual can become just as ill as if the disease were contracted under "real world" exposure conditions, although it is assumed that therapeutic countermeasures would be initiated as soon as infection was recognized. Data from such studies would also have to be considered in terms of inter-individual susceptibility.

3.5 Determining Compliance with Ethical Standards (Issue e)

Even if the Agency has ethical standards in place, there is the question of determining compliance with those standards. There is an imperative to actively oversee compliance with these standards on a continuing basis. This element of the Charge asks how the Agency can determine whether and to what extent its ethical standards have been met in a particular test with respect to a) informed consent, b) voluntary participation, and c) Institutional Review Boards.

Specifically, the Agency’s “having standards in place” means precisely the following: a) there is a policy describing the requirements for review and approval; and b) there is a mechanism for assuring compliance.

Attentive Agency oversight of compliance with its procedures for protection of human subjects requires written compliance oversight procedures. The procedures should be in sufficient detail so that researchers know what to expect, and, to that end, Agency procedures should be publicly promulgated and freely available. The Agency can expect, and should be prepared, to revise its compliance oversight procedures, as needed, to keep pace with evolving thinking and practice.

To pursue the goal of compliance oversight properly, the Agency will require staff with

the authority to carry out compliance oversight and to make formal determinations regarding noncompliance. As a matter of best practice, compliance oversight staff should be full-time individuals whose duties exclusively address compliance oversight. Individuals who are advocates for the rights and welfare of human subjects, who are committed to thoroughness, and who are unencumbered in their formulating and asking of pertinent questions should be selected for such a review staff.

Agency staff dedicated to compliance oversight should not be responsible for day-to-day education and interpretation of Agency standards regarding human-subject protection. It is critical to preserve an easy avenue for asking the Agency questions in a non-threatening atmosphere, and having those questions answered by Agency staff without nominal responsibility for "compliance oversight."

The following sections of this report discuss means by which the Agency can determine whether and to what extent its ethical standards have been met in a particular test, in the context of informed consent, voluntary participation, and IRBs.

3.5.1 Informed Consent

In reviewing proposed or submitted human studies, Agency staff should examine informed consent documents and informational brochures or allied materials, including advertisements intended to recruit subjects. "Advertisements" include electronic items posted on the World Wide Web. Agency staff should seek answers to the following questions of the informed consent document and process:

a) Are the required elements of information present?

b) Is the language understandable to the prospective subject?

c) Who actually seeks the consent of the subject?

3.5.2 Voluntary Participation

Agency staff should ask the following questions concerning the research under scrutiny:

a) What steps have been taken to minimize the possibility of coercion or undue influence?

b) How will the prospective subject be provided with sufficient opportunity to consider whether or not to participate?

c) What instruction is provided to research staff who will be recruiting subjects?

d) How many prospective subjects decline participation?

e) How many subjects withdraw from the research effort?

f) Is the design of the experiment valid? Has it sufficient power? Does it use the appropriate response measures?

3.5.3 Institutional Review Board (IRB)

As a matter of routine compliance oversight, Agency staff should a) validate membership of Institutional Review Boards, b) evaluate IRB policies and procedures, and c) review minutes of selected IRB meetings. All IRB records must be accessible for inspection and copying by authorized representatives of the Agency at reasonable times and in a reasonable manner (see 40 CFR Part 26.115(b)). These records must capture the identity of persons recruited for experimentation, including the total numbers, sex, ethnicity, and age. Given differing cultural and political systems, as well as the simple fact of distance, it is very difficult to maintain this level of scrutiny of foreign research activities. The Agency should consider it imperative to provide needed staff and financial resources to accomplish the best possible monitoring of foreign research whose results are presented to the Agency.

Agency staff should evaluate the IRB's receipt of reports of unanticipated problems involving risks to subjects or others. Agency staff should ask of the IRB, "What additional safeguards does the IRB require to protect the rights and welfare of subjects who are likely to be vulnerable to coercion or undue influence?"

There is no substitute for site visits in evaluating IRB compliance. The Agency should exercise, on occasion, authority to carry out "not-for-cause" on-site inspections and audits. Common knowledge of this Agency practice, despite the infrequency of such site visits, has a remarkable deterrent value. (This approach is similar in principle to the Internal Revenue Service random audit and its impact on compliance with income tax code.) The publicity that attends this Agency practice (i.e., the telling and retelling of stories of noncompliance) has broad positive impact for human research subjects.

Agency audits of the IRBs under its purview should include performance measures -- and not just the paper trail. The Agency should make certain that IRBs under its purview have sufficient provisions for meeting space and sufficient staff to support the IRB's review and record-keeping duties (see 40 CFR Part 26.103(b)(2)).

In short, compliance oversight requires an ongoing commitment on behalf of the Agency and its staff in the dynamic and evolving field of research ethics. This commitment must include the provision of sufficient staff and budget to maintain this oversight. Moreover, the Agency's effort would be well-served by creating an internal evaluation organization to facilitate oversight and maintain regular communication with other federal departments and agencies.

4. ADDITIONAL THOUGHTS

(NOTE TO THE SUBCOMMITTEE–The Co-chairs would like your opinion on the retention of this section. They are inclined to delete the entire section as contributing little to the overall "message" of the report. Please let us know your position on retaining or deleting the section.)

During the preparation of this report, certain suggestions and ideas not explicitly addressed during the public meeting were identified, particularly vis-vis the protection of human subjects.

As was discussed, EPA regulations at Title 40 Code of Federal Regulations Part 26 (40 CFR Part 26) have mandated protections for human subjects since 1991. Specifically, 40 CFR Part 26 applies to human-subject-research that is conducted, supported, or otherwise subject to regulation by the Agency (Sec. 26.101(a)). **Circumstances now dictate that the Agency take whatever administrative action is necessary to extend the protections of 40 CFR Part 26, as pertinent to those human-subject-research activities generating data that will be submitted to the Agency for review and consideration.** The Agency should move as soon as possible under the Administrative Procedures Act to promulgate an interim final rule to this effect. The interim final rule should have the effect that any submission to the Agency of data derived from human subjects will have undergone review and approval in compliance with Section 26.101, Section 102, and Section 26.107 through Section 26.117 prior to enrollment of any human subject.

With regard to certain groups of human subjects who may be especially vulnerable to coercion or undue influence, the Agency should immediately begin the rule-making process to formally adopt additional protections. The most expeditious and well-founded course for the Agency would be to adopt the additional protections that the U.S. Department of Health and Human Services has had in place for more than 15 years for pregnant women, prisoners, and children. In specific, the Agency should adopt, as part of 40 CFR Part 26, the additional protections for human subjects found at 45 CFR Part 46 Subpart B (pregnant women and fetuses), 45 CFR Part 46 Subpart C (prisoners), and 45 CFR Part 46 Subpart D (children). As it proceeds in this undertaking, the Agency, through the EPA Human Subjects Research Review Official, should consult and coordinate with the Human Subjects Research Subcommittee, Committee on Science, National Science and Technology Council.

The Agency should consider setting up its own review system. Under this scheme, registrants who anticipate collecting data from human research subjects would be obliged to seek prior

IRB review of the proposed human subject research and obtain approval from the EPA before proceeding to enroll human subjects in the research. **Additionally, the Agency may want to consider the need and feasibility of establishing human test guidelines for specific endpoints, similar to what has been established for various animal test endpoints and human exposure studies.**

5 MAJOR RECOMMENDATIONS

In the body of this of this report, the Subcommittee has provided (within the context of the Charge) many recommendations and cautions to the EPA. This section "looks across" the Charge and highlights the Subcommittee's major findings and advice to the Agency. These findings are:

a) All research involving humans should require prior review by an Institutional Review Board.

b) The structure, function, and activities of both the Agency's IRBs and external IRBs of entities submitting data should be under active and aggressive scrutiny by EPA, with adequate staff and financial resources provided to carry out this mission. EPA should establish an internal ethics review organization to perform this function, staffed by full-time individuals whose duties address exclusively compliance oversight. The review organization should also provide an institutional focus for continuous close liaison on ethical matters with other federal agencies.

c) The Subcommittee believes that intentional administration of pesticides to human subjects testing is acceptable, subject to limitations ranging from "rigorous" to "severe." Those supporting such testing feel that the information sought must not be available via other sources (e.g., animal studies and models or study of incidental exposures), and the information expected to be gained must promise reasonable benefits to the individual or society at large. Studies should be appropriately designed to address the stated objective, and have sufficient statistical power to provide an unambiguous answer to the question under investigation.

d) In no case should developing humans (i.e., the fetus, infant, and young children) be exposed to neurotoxic chemicals. There are currently too many unknown dangers to justify such studies, even under the most extraordinary circumstances.

e) The EPA should take whatever administrative action is necessary to extend the protections of 40 CFR Part 26 to all human research activities whose results will be submitted to the Agency.

f) Some of the Subcommittee Members that accepted the use of human volunteer testing of pesticides identified certain situations in which such testing would or would not be appropriate:

1) It would not be appropriate to conduct human volunteer testing when adequate human data are already available.

2) Human volunteer studies would not be appropriate for pesticides in use today when data of equal quality can be obtained from field exposure studies.

3) Human volunteer studies could be appropriate when there are significant data gaps and such studies would provide a more accurate risk assessment.

4) Human volunteer studies could be appropriate for pesticides, which are not yet on the market, i.e. new pesticides.

5) Given the significance of statistical considerations in regard to human study design, we encourage the Agency to organize a workshop to specifically deal with this issue.

APPENDIX A - FACTORS AFFECTING STATISTICAL POWER

Major factors are:

a) Clinical trials for pharmaceuticals fall into two basic areas: evaluation of dose ranges for proper delivery of the agent, and larger studies aimed at product efficacy. With pesticides, controlled experimental exposures are aimed at similar issues, comprising absorption, distribution, metabolism, and elimination (ADME) studies and studies aimed at finding exposures intended (and expected) to produce some trivial, non-toxic effect in the study subjects (NOAEL). For ADME studies, one is attempting to estimate pharmacokinetic-pharmacodynamic parameters. The precision required for the estimation of these parameters is determined by knowledge of the variability in the general population and by a decision about the size of the standard error relative to the mean value of the population. Generally, the ratio of the standard error to the mean should be smaller than 1.0, preferentially much smaller than 1.0. In attempting to find a dose that produces effects no larger than a specified value, the probability that the effect is greater than the specified value should be fairly small, typically less than 0.2 or 0.1. Different designs can satisfy these requirements and care should be taken to have the design match the needs. Similar concepts apply to human epidemiology studies and human studies of biomarkers in worker or environmentally exposed populations.

b) Questions about the precision of estimates and the probability of exceedance should also be addressed. Have statistical criteria been established to allow for continuous monitoring of the responses in such a way that, if the question can be answered earlier than projected, the study is terminated? Statistical methods exist for evaluating these issues without affecting the final probability of making an error. Sequential decision designs, such as those now recommended for LD_{50} calculations, could also serve such a purpose. In fact, they could also be applied in short-term experiments.

Several Members of the Subcommittee expressed serious reservations concerning the overall issue of statistical considerations in regard to human study design. Some of these Members felt that this issue was of such import that it deserved separate consideration. Therefore, we encourage the Agency to organize a workshop to specifically deal with this issue.

APPENDIX B – STATISTICAL CONSIDERATIONS IN NOAEL STUDIES

Two critical statistical measures determine the ability of a study to meet its objectives: the probability of detecting an effect when no effect exists (Type I or alpha error); and the probability of missing an effect that is real (Type II or beta error). The probability of detecting a true effect is generally referred to as the power and is defined as 1 minus beta. The NOAEL is partially defined by rejection of the null hypothesis (i.e., that no effect exists).[2]

The choice of an effect size to look for in a study of a neurotoxic pesticide is somewhat arbitrary and entails value judgments. For example, what effect size should be sought in a study of dietary pesticides? The number of exposed American children argues for conducting studies that will find small effect sizes. In this context the word "small" does not mean negligible; it means difficult to measure. There are 18.9 million children under five years of age in the United States. If a pesticide in their diet and environment were to cause a 1% increase in the rate of neurobehavioral toxicity, that would be 189,000 affected cases. Assuming a base rate of deficit of 1%, we can ask how many subjects would be needed to find an increase from 1% to 2%, or from 3% to 4%. The proportion of children 3- 5 with disabilities is approximately 4%. We also calculate the number of subjects required to find an increase from 4 to 5% with an 80% power.

Alpha level	**.05**	**.05**	**.05**	**.05**
proportion in unexposed group	**.01**	**.02**	**.03**	**.04**
proportion in exposed group	**.02**	**.03**	**.04**	**.05**
number of cases in each group	**3017**	**5071**	**7062**	**6725**
POWER	**.90**	**.90**	**.90**	**.80**

Entering the number of subjects commonly used in past human studies made available to the EPA enables us to measure the power to find an adverse pesticide effect

Alpha level	**.05**	**.05**	**.05**
proportion in unexposed group	**.04**	**.04**	**.04**
proportion in exposed group	**.05**	**.05**	**.05**

[2]The alpha level is generally specified in advance of the study. The beta error, and therefore the power of a study, is determined by three factors: the alpha level initially set, the size of the effect looked for, and the number of subjects studied. If any three parameters are established, the fourth is fixed and readily determined. If the effect size sought, the alpha level and the power desired are known, the number of subjects can easily be calculated. If the alpha level, effect size and number of subjects are known, the power can be determined.

number of cases in each group	10	20	50
POWER	.03	.04	.04

It can be readily seen that large numbers of subjects (between 6000 and 14,000) are needed to make a dependable no-effect assertion for a small effect with 80% confidence. Conversely, with the number of subjects employed by registrants in past studies submitted to EPA, there was little chance of finding an effect if it were present. A power of 0.04 is one chance in 25. It is as if there were 4 black balls representing a toxic effect and 96 white balls representing no toxic effect placed in a jar. Asserting that no toxicity was seen in a study of 50 subjects is no different that reaching into the jar, pulling out a white ball, and stating that only white balls were in the jar.

So, what is the probability of missing a real effect for a given sample size and a given true effect? To be able to study this issue, one must know the distribution of the target measurement in the study population and have some idea of how changes in this value will affect this distribution.

To illustrate the value of power for continuous alternatives, consider the levels of acetylcholinesterase in humans. Singh *et al*. (1987) measured acetylcholinesterase (nmol/mg HB/min) in the red blood cells of 193 individuals in India who were "unexposed" to organophosphates. They estimated a mean of 35 and a standard deviation of 13.7. Assuming the variance acetylcholinesterase (AchE) in this population is due to two independent sources of variation, variation across individuals (50%) and variation within individuals (50%), and assuming the reduction in AchE is still subject to interindividual variation and a small additional variation due to variation in response to the organophosphate, one can estimate the power for detecting a real effect for various reductions in AchE levels and various sample sizes (see code below for the parameters used to make these calculations). Table 1 presents the power of the signed-rank test for AchE reduction in the case where individuals are used as their own controls and comparisons are made between a targeted time point with the specified reduction and the AchE level prior to exposure. It is clear that, if the sample size is greater than 10, it is possible to detect a 25% or greater reduction in AChE with high power. However, for a 10% reduction, at least 20 samples must be taken, for a 10% reduction, at least 100 samples must be taken and for a 1% reduction, at least 1000 samples must be taken.

It is possible to argue that since we have used NOAEL's from animal studies as a general rule for setting standards, then the power for the animal study should equal the power of the human study in detecting a NOAEL. All else equal, this would mean equivalent samples. If there are differences in variation between the species, the sample sizes would have to be adjusted. Even if the powers for detecting a NOAEL are equivalent, it should be noted that the human study will

provide less protection against a possible adverse effect since the 10-fold interspecies extrapolation uncertainty factor will not be applied.

The proper way to design a human study would be to decide upon a change in AchE levels which would be of no clinical significance taking into account sensitive individuals and possible effects of longer exposures in the environment as compared to the laboratory. Then choose exposures which are unlikely to yield this level of response and choose a sample size such that, if this response were true, you would have sufficient power to detect it. Even this approach carries some risk since some members of the study population could be somewhat sensitive to the exposure. In general, the targeted reduction should be fairly low to insure safety (say less than 5% or less than 1%). This would require sample sizes much larger than those generally used in these types of trials.

TABLE 3: Power (in %) for detecting specified changes in AchE levels based upon distributions and assumptions given in MatLab code following the table

Sample Size	Reduction In AchE (AchE in nmol/mg HB/min				
	50%	25%	10%	5%	1%
10	1	99.6	56.2	15.2	5.0
20	1	1	89.6	35.8	7.6
50	1	1	99.8	76.2	8.2
100	1	1	1	96.6	11.6
200	1	1	1	99.8	18.8
500	1	1	1	1	41.0
1000	1	1	1	1	69.8
2000	1	1	1	1	94.0

(NOTE TO THE SUBCOMMITTEE–The Co-chairs would like your opinion on the retention of this section. They are inclined to delete the entire section as not contributing to the overall "message" of the report. Please let us know your position on retaining or deleting the section.)

APPENDIX C - ETHICAL RATINGS FOR SAMPLE HUMAN EXPOSURE SCENARIOS

To capture the sense of the Committee's deliberations, the following scenarios were synthesized to exemplify some of the ethical issues before it. Although the examples were not drawn directly from the available literature, they embody relevant approaches and questions in those publications. For each scenario, a structured review might include two phases. The first phase, "Critique," would come in the form of an assessment in which features such as statistical power, scientific validity, informed consent and the adequacy of medical monitoring would be examined. The second phase, "Ratings," is a more quantitative exercise. As shown in Table 1, each member of a selected panel would provide a judgment of which of the 25 cells in the matrix describes the proposed study. One dimension of the matrix corresponds to Health Risks and the other to what might be termed Ethics Risks. Although it may prove difficult to separate the two dimensions, it is a useful exercise to make such an attempt. For example, a protocol viewed as offering a negligible health risk might be acceptable with a moderate ethical risk.

Five sample scenarios (PestTest A-E) are given below. The Committee recomends that the Agency compile a set of hypothetical scenarios such as these, along with their associated Critiques and Health/Ethics ratings, to guide policy personnel. Sample scenarios and associated critiques and ratings would also help instruct those designing volunteer studies about Agency standards for acceptable protocols and data.

<u>PestTest A</u>

In a rat study, the major urinary metabolite (50%) of this compound after oral administration was identified as META1. After dermal application, the major metabolite was identified as META2 (40%), with META1 falling to 16%. Urinary samples from field workers, exposed predominantly by skin contact, showed META1 to comprise 40% of the urinary metabolites. Because of the discrepancies with the rat data, a test in volunteers is now proposed to study metabolite profiles after both oral and dermal exposure routes, and to calculate various pharmacokinetic parameters. The oral dose corresponds to the RfD. Dermal exposure will require one forearm placed in the commercial formulation for 20 minutes. The proposed sample size is eight, and the subjects, all male, will be drawn from advertisements posted in a college campus newspaper. After dosing, the subject will be observed for 24 hours in the clinical facility.

Critique:

Ratings:

Health Risk:

Ethics Risk:

PestTest B

This insecticide is typically sprayed on crops and is dissolved in an organic solvent. Because field conditions are highly variable, a study is now proposed to examine the relationship between ambient air concentrations and blood, urine, and expired air concentrations. Exposure will take place in a human inhalation facility. The concentrations to be used are about 10% of the mean measured under field conditions. The proposed sample size is 18 young, healthy males. They will be exposed for one hour and followed for eight hours in the facility.

Critique:

Ratings:

Health Risk:

Ethics Risk:

PestTest C

Several reports indicate residual neuropsychological deficits following clinical recovery from occupational exposure to organophosphate pesticides at levels high enough to induce clinical signs. One way to trace such residual neurological effects might be to obtain sensitive electrophysiological measures. Checkerboard contrast-reversal visual evoked potentials (VEPs) from scalp electrodes will be obtained from young male volunteers exposed, by oral dosing, to a dose high enough to produce a reduction in erythrocyte AChE to 50% of normal. The primary measures will be VEP latency and amplitude. A sample of 10 subjects will first be studied to estimate inter-individual variability in these measures; the data will serve to select an appropriate sample size, with power specified as 80% at p=.05 The electrophysiological indices will be measured at 4 and 24 hours after dosing and weekly thereafter for 6 weeks. AChE will be assayed at the same time points.

Critique:

Ratings:

Health Risk:

Ethics Risk:

PestTest D

Serum paraoxonase concentrations influence the metabolism of organophosphate pesticides in humans. Two different PON1 genotypes, with significantly different detoxifying capability, will be selected from a population of agricultural workers. Notices will be posted at hiring centers offering fees of $50 for providing blood samples for PON1 analysis. Offers will then be made to members of the two groups for participation in a study to establish the pharmacokinetic profile of a newly registered OP insecticide. The fee will be $150. A total of 30 subjects per group will be recruited. The dose chosen is equivalent to one-tenth the NOAEL based on rat data.

Critique:

Ratings:
Health Risk:
Ethics Risk:

<u>PestTest E</u>

The RfD of this compound, based on chronic rat data (NOAEL/100) is 1.0 mg/kg. A study is proposed to administer the compound orally to young male adults for 30 days at a dose of 0.1 mg/kg, which would incorporate the additional 10-fold uncertainty factor of the FQPA. A total of 20 subjects is projected. The primary endpoint will be erythrocyte AChE, accompanied by weekly clinical medical examinations. The subjects will ingest one capsule containing the compound daily.

Critique:

Rating:
Health Risk:
Ethics Risk:

Table 1. A sample of raters, such as experienced EPA risk assessors, or a group of bioethics specialists, or a sample of occupational physicians, etc., is given the proposed protocol. Each rater then marks the cell appropriate to his or her judgment of where the health and ethics risks lie. Summary statistics for a panel might be expressed, for instance, as Box Plots. EPA policy would dictate the region of the matrix that is acceptable for a particular class of pesticides. For example, the Agency might decide to restrict acceptable human studies to cells 1 and 2 on both axes, as shown in the gray areas of the matrix.

	ETHICS RISK				
HEALTH RISK	**1**	**2**	**3**	**4**	**5**
5					
4					
3					
2					
1					

Health Risk Ratings:

1 No adverse effects
2 Mild acute discomfort
3 Moderately aversive, acute
4 Highly aversive, acute
5 Prolonged neurotoxicity

Ethics Risk Ratings:

1 No ethical problems
2 Minor ethical problems, e.g., advertisement
3 Questionable procedures, e.g., excessive compensation
4 Major ethical problems, e.g., misleading consent form
5 IRB rejection recommended

REFERENCES

ACHRE (Advisory Committee on Human Radiation Experiments). 1996a. Research ethics and the medical profession: report of the advisory committee on human radiation experiments. Journal of the American Medical Association, 276:403-409.

ACHRE (Advisory Committee on Human Radiation Experiments). 1996b. The Human Radiation Experiments. Oxford University Press, New York.

ACHRE (Advisory Committee on Human Radiation Experiments), Final Report, October 1995, Washington DC.

Beecher, H. K. 1966. Ethics and Clinical Research. New England Journal of Medicine 274.

Bois, F.Y., Smith, T.S., Gelman, A., Chang, A.H-T., and A.E. Smith. 1999. Optimal design for a study of butadiene toxicokinetics in humans. *Toxicological Sciences*, 49:213-224.

Bond, J.A., Himmelstein, M.W., and M.A. Medinski. 1996. The use of toxicological data in mechanistic risk assessment: 1,3-butadiene as a case example. *Int. Archiv. of Occ. and Env. Health*, 68:415-420.

Bureau of the Census. 1993. http://www.census.gov/hhes/www/saipe/saipe93/estimate.html

Caplan, A. 1998. Am I My Brother's Keeper?: Ethical Issues on the Frontier of Biomedicine, Indiana University Press.

Caplan, A. 1993. How Should Science Deal with Data from Unethical Research. *The Journal of NIH Research*, 5, 5, pp. 22-26.

Caplan, A. 1992. When Evil Intrudes: The Use of Scientific Information Acquired from the Tuskegee Syphilis Study," Hastings Center Report, Vol 22, pp. 29-32.

Caplan, A. 1992. When Medicine Went Mad, Bioethics and the Holocaust. Humana Press.

Cherian MG, Hursh JB, Clarkson TW, Allen J. Radioactive mercury distribution in biological fluids and excretion in human subjects after inhalation of mercury vapor. Arch Environ Health 1978;33:109-114.

Iregren, A. 1966. Behavioral methods and organic solvents: questions and consequences. *Environ Health Perspect*,104, Suppl 2, 361-366

EPA (Environmental Protection Agency). 19**9X**. Neurotoxicolgical Assessment Guidelines

Jones, J.H. 1981. Bad Blood. The Free Press, New York, NY.

Mantel, XX. and XX. Bryan. 1961. Safety testing of carcinogenic agents, J*NCI* 27: 455-470.

(NTP) National Toxicology Program. 1993. Toxicological and carcinogenetic studies of 1,3-butadiene. NTP Technical Report Series No. 434. Research Triangle Park, NC.

SAB. 19XX. AChe

SAB. 19XX. AChe 2

Stevens, R., Spurgeon, A., and H. Berry. 1996. Organophosphates: the relationship between chronic and acute exposure effects. Neurotoxicol. Teratol. 18: 449-453.

Steenland, K., Jenkins, B., Ames, R.G., O'Malley, M, Chrislip, D., and J. Russo. 1994. Chronic neurological sequelae to organophosphate pesticide poisoning. Am J Public Health May; 84(5):731-6.

Stern, P.C., and H.V. Fineberg. (eds). 1996. Understanding Risk: Informing Decisions in a Democratic Society. Washington, DC, National Academy Press.